

# Weakly Submodular Functions

Allan Borodin\*

Dai Tri Man Lê†

Yuli Ye‡

July 31, 2014

## Abstract

Submodular functions are well-studied in combinatorial optimization, game theory and economics. The natural diminishing returns property makes them suitable for many applications. We study an extension of monotone submodular functions, which we call *weakly submodular functions*. Our extension is somewhat unusual in that it includes some (mildly) supermodular functions. We show that several natural functions belong to this class.

We consider the optimization problem of maximizing a weakly submodular function subject to uniform and general matroid constraints. For a uniform matroid constraint, the “standard greedy algorithm” achieves a constant approximation ratio where the constant (experimentally) converges to 5.95 as the cardinality constraint increases. For a general matroid constraint, a simple local search algorithm achieves a constant approximation ratio where the constant (analytically) converges to 10.22 as the rank of the matroid increases.

## 1 Introduction

There are many applications where the goal becomes a problem of maximizing a submodular function subject to some constraint. In many applications the submodular function  $f$  is also monotone, non-negative and normalized so that  $f(\emptyset) = 0$ . Such applications arise for example in the consideration of influence in a stochastic social network as formalized in Kempe, Kleinberg and Tardos [9], diversified search ranking as in Bansal, Jain, Kazeykina and Naor [3] and in document summarization as in Lin and Bilmes [12]. In another application, following Gollapudi and Sharma [8], Borodin, Lee and Ye [5] considered the linear combination of a monotone submodular function that measures the “quality” of a set of results combined with a diversity function given by the max-sum dispersion measure, a widely studied measure of diversity. Their analysis suggested that although the max-sum dispersion measure is a supermodular function, it possessed similar properties to monotone submodular functions. In this paper we develop this idea by introducing the class of weakly submodular functions and show that greedy and local search algorithms can be used (respectively) to maximize such functions subject to a cardinality (resp. matroid) constraint.

## 2 Preliminaries

Let  $f$  be a set function over a universe  $U$  satisfying the following properties:

---

\*Dept of Computer Science, University of Toronto, bor@cs.toronto.edu

†Dept of Computer Science, University of Toronto, ledt@cs.toronto.edu

‡Amazon.com, Inc, yeyuli@amazon.com

- $f(\emptyset) = 0$ ; i.e.  $f$  is normalized.
- $f(S) \geq 0$  for all  $S \subseteq U$ ; i.e.  $f$  is non-negative
- $f(S) \leq f(T)$  for all  $S \subseteq T \subseteq U$ ; i.e.  $f$  is monotone

A function  $f(\cdot)$  is submodular if for any two sets  $S$  and  $T$ , we have

$$f(S) + f(T) \geq f(S \cup T) + f(S \cap T).$$

We define the following generalization. We call a function  $f(\cdot)$  *weakly submodular* if for any two sets  $S$  and  $T$ , we have

$$|T|f(S) + |S|f(T) \geq |S \cap T|f(S \cup T) + |S \cup T|f(S \cap T).$$

### 3 Examples of Weakly Submodular Functions

There are several natural examples of weakly submodular functions. Our examples of weakly submodular functions are all normalized, non-negative and monotone.

#### 3.1 Submodular Functions

From the weakly submodular definition, it is not obvious that monotone submodular functions are a subclass of weakly submodular functions. We will prove that this is indeed the case.

**Proposition 3.1** *Any monotone submodular function is weakly submodular. This, of course, implies that every linear function is a weakly submodular.*

**Proof:** Given a monotone submodular function  $f(\cdot)$  and two subsets  $S$  and  $T$ , without loss of generality, we assume  $|S| \leq |T|$ , then

$$|T|f(S) + |S|f(T) = |S|[f(S) + f(T)] + (|T| - |S|)f(S).$$

By submodularity  $f(S) + f(T) \geq f(T \cup S) + f(T \cap S)$  and monotonicity  $f(S) \geq f(S \cap T)$ , we have

$$\begin{aligned} |T|f(S) + |S|f(T) &= |S|[f(S) + f(T)] + (|T| - |S|)f(S) \\ &\geq |S|[f(S \cup T) + f(S \cap T)] + (|T| - |S|)f(S \cap T) \\ &= |S|f(S \cup T) + |T|f(S \cap T) \\ &= |S \cap T|f(S \cup T) + [(|S| - |S \cap T|)f(S \cup T) + |T|f(S \cap T)]. \end{aligned}$$

And again by monotonicity  $f(S \cup T) \geq f(S \cap T)$ , we have

$$(|S| - |S \cap T|)f(S \cup T) + |T|f(S \cap T) \geq (|S| + |T| - |S \cap T|)f(S \cap T) = |S \cup T|f(S \cap T).$$

Therefore

$$|T|f(S) + |S|f(T) \geq |S \cap T|f(S \cup T) + |S \cup T|f(S \cap T);$$

the proposition follows.  $\square$

We note that the proof of Proposition 3.1 did not require the function  $f(\cdot)$  be normalized or non-negative. But the proof did use the monotonicity of  $f(\cdot)$ . Non-monotone submodular functions (such a Max-Cut and Max-Di-Cut) are also widely studied. In contrast to Proposition 3.1, if we extend the weakly submodular definition to non-monotone functions, then it is no longer the case that a non-monotone submodular function would necessarily be a non-monotone weakly submodular function.

**Proposition 3.2** *There is a non-monotone submodular function  $f(\cdot)$  that is not weakly submodular. More specifically, the Max-Cut function (for a particular graph  $G$ ) is not weakly submodular.*

**Proof:** Consider a graph  $G = (U, E)$  where  $V = R \cup \{s\} \cup \{t\}$  and  $E = \{(s, u), (u, t) | u \in R\}$ . Letting  $S = R \cup \{s\}$  and  $T = R \cup \{t\}$ , we have the following letting for  $|R| = n$ .

- $f(S) = f(T) = n$
- $f(S \cup T) = f(U) = 0$
- $f(S \cap T) = f(R) = 2n$

We have

1.  $|T|f(S) + |S|f(T) = (n+1)n + (n+1)n = 2n^2 + 2n$
2.  $|S \cap T|f(S \cup T) + |S \cup T|f(S \cap T) = n \cdot 0 + (n+2) \cdot 2n = 2n^2 + 4n$

This contradicts the weakly submodular definition.  $\square$

Hereafter, we will we restrict attention to monotone, non-negative and normalized, functions.

### 3.2 Sum of Metric Distances of a Set

Let  $U$  be a metric space with a distance function  $d(\cdot, \cdot)$ . For any subset  $S$ , define  $d(S)$  to be the sum of distances induced by  $S$ ; i.e.,

$$d(S) = \sum_{\{u, v\} \subseteq S} d(u, v)$$

where  $d(u, v)$  measures the distance between  $u$  and  $v$ . The problem of maximizing  $d(S)$  (subject to say a cardinality or matroid constraint) is one of many dispersion problems studied in location theory.

We also extend the function to a pair of disjoint subsets  $S$  and  $T$  and define  $d(S, T)$  to be the sum of distances between  $S$  and  $T$ ; i.e.,

$$d(S, T) = \sum_{u \in S, v \in T} d(u, v).$$

We have the following proposition.

**Proposition 3.3** *The sum of metric distances  $d(S)$  of a set is weakly submodular (and clearly monotone).*

**Proof:** Given two subsets  $S$  and  $T$  of  $U$ , let  $A = S \setminus T$ ,  $B = T \setminus S$  and  $C = S \cap T$ . Observe the fact that by the triangle inequality, we have

$$|B|d(A, C) + |A|d(B, C) \geq |C|d(A, B).$$

Therefore,

$$\begin{aligned}
& |T|d(S) + |S|d(T) \\
&= (|B| + |C|)[d(A) + d(C) + d(A, C)] + (|A| + |C|)[d(B) + d(C) + d(B, C)] \\
&= |C|[d(A) + d(B) + d(C) + d(A, C) + d(B, C)] + (|A| + |B| + |C|)d(C) \\
&\quad + |B|d(A) + |A|d(B) + |B|d(A, C) + |A|d(B, C) \\
&\geq |C|[d(A) + d(B) + d(C) + d(A, C) + d(B, C)] + |S \cup T|d(S \cap T) + |C|d(A, B) \\
&= |C|[d(A) + d(B) + d(C) + d(A, C) + d(B, C) + d(A, B)] + |S \cup T|d(S \cap T) \\
&= |S \cap T|d(S \cup T) + |S \cup T|d(S \cap T).
\end{aligned}$$

□

### 3.3 Average Non-Negative Segmentation Functions

Motivated by applications in clustering and data mining, Kleinberg, Papadimitriou and Raghavan [10] introduce the general class of segmentation functions. In their generality, segmentation functions need not be submodular nor monotone. They show that every segmentation belongs to call they call *meta-submodular functions* and consider the greedy algorithm for “weakly monotone” meta-submodular functions. We now consider another broad class of segmentation functions.

Given an  $m \times n$  matrix  $M$  and any subset  $S \subseteq [m]$ , a *segmentation function*  $\sigma(S)$  is the sum of the maximum elements of each column whose row indices appear in  $S$ ; i.e.;  $\sigma(S) = \sum_{j=1}^n \max_{i \in S} M_{ij}$ . A segmentation function is *average non-negative* if for each row  $i$ , the sum of all entries of  $M$  is non-negative; i.e.,  $\sum_{j=1}^n M_{ij} \geq 0$ .

We can use columns to model individuals, and rows to model items, then each entry of  $M_{ij}$  represents how much the individual  $j$  likes the item  $i$ . The average non-negative property basically requires that for each item  $i$ , on average people do not hate it. Next, we show that an average non-negative segmentation function is weakly-submodular. We first prove the following two lemmas.

**Lemma 3.4** *An average non-negative segmentation function is monotone.*

**Proof:** Let  $S$  be a proper subset of  $[m]$ , and  $e$  be an element in  $[m]$  that is not in  $S$ . If  $S$  is empty, then by the average non-negative property, we have  $\sigma(\{e\}) = \sum_{j=1}^n M_{ej} \geq 0$ . Otherwise, by adding  $e$  to  $S$  we have  $\max_{i \in S \cup \{e\}} M_{ij} \geq \max_{i \in S} M_{ij}$  for all  $1 \leq j \leq n$ . Therefore  $\sigma(S \cup \{e\}) \geq \sigma(S)$ . □

**Lemma 3.5** *For any non-disjoint set  $S$  and  $T$  and an average non-negative segmentation function  $\sigma(\cdot)$ , we have*

$$\sigma(S) + \sigma(T) \geq \sigma(S \cup T) + \sigma(S \cap T).$$

*This is also referred as the meta-submodular property [11].*

**Proof:** For any non-disjoint set  $S$  and  $T$  and an average non-negative segmentation function  $\sigma(\cdot)$ , we let  $\sigma_j(S) = \max_{i \in S} M_{ij}$ . We show a stronger statement that for any  $j \in [n]$ , we have

$$\sigma_j(S) + \sigma_j(T) \geq \sigma_j(S \cup T) + \sigma_j(S \cap T).$$

Let  $e$  be an element in  $S \cup T$  such that  $M_{ej}$  is maximum. Without loss of generality, assume  $e \in S$ , then  $\sigma_j(S) = \sigma_j(S \cup T) = M_{ej}$ . Since  $S \cap T \subseteq T$ , we have  $\sigma_j(T) \geq \sigma_j(S \cap T)$ . Therefore,

$$\sigma_j(S) + \sigma_j(T) \geq \sigma_j(S \cup T) + \sigma_j(S \cap T).$$

Summing over all  $j \in [n]$ , we have

$$\sigma(S) + \sigma(T) \geq \sigma(S \cup T) + \sigma(S \cap T)$$

as desired.  $\square$

**Proposition 3.6** *Any average non-negative segmentation function is weakly submodular.*

**Proof:** For any two set  $S$  and  $T$  and an average non-negative segmentation function  $\sigma(\cdot)$ , if  $S$  and  $T$  are non-disjoint then by Lemma 3.5,  $S$  and  $T$  satisfy the submodular property and hence they satisfy the weakly submodular property by Proposition 3.1. If  $S$  and  $T$  are disjoint, then  $|S \cap T| = 0$ , and  $|S \cup T| = |S| + |T|$ . By monotonicity property in Lemma 3.3, we also have  $\sigma(S) \geq \sigma(S \cap T)$  and  $\sigma(T) \geq \sigma(S \cap T)$ . Therefore,

$$|S \cap T|\sigma(S \cup T) + |S \cup T|\sigma(S \cap T) \leq |T|\sigma(S \cap T) + |S|\sigma(S \cap T) \leq |T|\sigma(S) + |S|\sigma(T);$$

the weakly submodular property is also satisfied.  $\square$

### 3.4 Small Powers of the Cardinality of a Set

Clearly, for any positive integer  $k$ , the functions  $f(S) = |S|^k$  can be computed in time  $O(\log k)$ . However, given Lemma 3.10 below, it is still useful to know what simple functions can be used in conjunction with other submodular and weakly submodular functions.

It is immediate to see that the functions  $f(S) = |S|^0$  and  $f(S) = |S|^1$  are linear and hence submodular. We will show that the square and the cube of the cardinality of a set are also weakly submodular.

**Proposition 3.7** *The square of cardinality of a set is weakly submodular.*

**Proof:** Given two subsets  $S$  and  $T$  of  $U$ , let  $a = |S \setminus T|$ ,  $b = |T \setminus S|$  and  $c = |S \cap T|$ .

$$\begin{aligned} & |T|f(S) + |S|f(T) \\ &= (b+c)(a+c)^2 + (a+c)(b+c)^2 \\ &= (a+b+2c)(b+c)(a+c) \\ &= (a+b+2c)(ab+ac+bc+c^2) \\ &\geq (a+b+2c)(ac+bc+c^2) \\ &= (a+b+2c)c(a+b+c) \\ &= c(a+b+c)^2 + (a+b+c)c^2 \\ &= |S \cap T|f(S \cup T) + |S \cup T|f(S \cap T). \end{aligned}$$

$\square$

**Proposition 3.8** *The cube of cardinality of a set is weakly submodular.*

**Proof:** Given two subsets  $S$  and  $T$  of  $U$ , let  $a = |S \setminus T|$ ,  $b = |T \setminus S|$  and  $c = |S \cap T|$ .

$$\begin{aligned}
& |T|f(S) + |S|f(T) \\
&= (b+c)(a+c)^3 + (a+c)(b+c)^3 \\
&= (a^2 + b^2 + 2c^2 + 2ac + 2bc)(b+c)(a+c) \\
&= [(a+b+c)^2 + c^2 - 2ab][ab + c(a+b+c)] \\
&= [(a+b+c)^2 + c^2][c(a+b+c)] + ab[(a+b+c)^2 + c^2] - 2a^2b^2 - 2abc(a+b+c) \\
&= c(a+b+c)^3 + c^3(a+b+c) + ab[(a+b+c)^2 + c^2 - 2ab - 2c(a+b+c)] \\
&= |S \cap T|f(S \cup T) + |S \cup T|f(S \cap T) + ab(a^2 + b^2 + c^2 + 2ab + 2ac + 2bc + c^2 - 2ab - 2ac - 2bc - 2c^2) \\
&= |S \cap T|f(S \cup T) + |S \cup T|f(S \cap T) + ab(a^2 + b^2) \\
&\geq |S \cap T|f(S \cup T) + |S \cup T|f(S \cap T).
\end{aligned}$$

□

It is easy to see that the function is weakly submodular for  $f(S) = |S|^0$  and  $f(S) = |S|^1$ . We now give an example that shows  $f(S) = |S|^4$  is not weakly submodular.

### 3.4.1 Higher powers

**Proposition 3.9**  $f(S) = |S|^4$  is not weakly submodular.

**Proof:** Given two subsets  $S$  and  $T$  of  $U$ , let  $a = |S \setminus T|$ ,  $b = |T \setminus S|$  and  $c = |S \cap T|$ . Suppose  $a = 4, b = 4, c = 1$ .

$$|T|f(S) + |S|f(T) = (b+c)(a+c)^4 + (a+c)(b+c)^4 = 6250$$

On the other hand, we have

$$|S \cap T|f(S \cup T) + |S \cup T|f(S \cap T) = c(a+b+c)^4 + c^4(a+b+c) = 9^4 + 9 = 6570$$

Therefore, the function is not weakly submodular.

Similarly, one can see that  $f(S) = |S|^k$  is not weakly submodular for all integers  $k \geq 4$ .

## 3.5 Linear combinations of weakly submodular functions

Next we show a basic but important property of weakly submodular functions.

**Lemma 3.10** *Non-negative linear combinations of weakly submodular functions are weakly submodular.*

**Proof:** Consider weakly submodular functions  $f_1, f_2, \dots, f_n$  and non-negative numbers  $\alpha_1, \alpha_2, \dots, \alpha_n$ . Let  $g(S) = \sum_{i=1}^n \alpha_i f_i(S)$ , then for any two set  $S$  and  $T$ , we have

$$\begin{aligned}
& |T|g(S) + |S|g(T) \\
&= |T| \sum_{i=1}^n \alpha_i f_i(S) + |S| \sum_{i=1}^n \alpha_i f_i(T) \\
&= \sum_{i=1}^n \alpha_i [|T|f_i(S) + |S|f_i(T)] \\
&\geq \sum_{i=1}^n \alpha_i [|S \cap T|f_i(S \cup T) + |S \cup T|f_i(S \cap T)] \\
&= |S \cap T| \sum_{i=1}^n \alpha_i f_i(S \cup T) + |S \cup T| \sum_{i=1}^n \alpha_i f_i(S \cap T) \\
&= |S \cap T|g(S \cup T) + |S \cup T|g(S \cap T).
\end{aligned}$$

Therefore,  $g(S)$  is weakly submodular.  $\square$

We now show two more examples of weakly submodular function using Lemma 3.10.

### 3.6 The Objective Function of Max-Sum Diversification

**Corollary 3.11** *The objective function of the max-sum diversification problem is weakly submodular.*

**Proof:** This follows immediate from Proposition 3.1 and 3.3 and Lemma 3.10.  $\square$

### 3.7 Restricted Polynomial Function on the Cardinality of a Set

**Corollary 3.12** *For polynomial function on the cardinality of a set, if the degree is less than four and coefficients are all non-negative, then the function is weakly submodular.*

**Proof:** This follows immediate from Proposition 3.7 and 3.8 and Lemma 3.10.  $\square$

## 4 Weakly Submodular Function Maximization Subject to a Cardinality Constraint

We emphasize again that we restrict attention to monotone, non-negative and normalized functions. In this section, we discuss a greedy approximation algorithm for maximizing weakly submodular functions subject to a uniform matroid (i.e cardinality constraint). In section 5 we consider an arbitrary matroid constraint.

Given an underlying set  $U$  and a weakly submodular function  $f(\cdot)$  defined on every subset of  $U$ , the goal is to select a subset  $S$  maximizing  $f(S)$  subject to a cardinality constraint  $|S| \leq p$ . We consider the following standard greedy algorithm that achieves approximation ratio  $\frac{e}{e-1}$  for monotone submodular maximization by a classic result of Nemhauser, Fisher and Wolsey [13]. Furthermore, they showed that this is the best approximation possible in the value oracle model and Feige [7] showed the same inapproximation holds for an explicitly defined function subject to the conjecture that  $RP \neq NP$ .

GREEDY ALGORITHM FOR WEAKLY SUBMODULAR FUNCTION MAXIMIZATION

```

1:  $S = \emptyset$ 
2: while  $|S| < p$  do
3:   Find  $u \in U \setminus S$  maximizing  $f(S \cup \{u\}) - f(S)$ 
4:    $S = S \cup \{u\}$ 
5: end while
6: return  $S$ 

```

**Theorem 4.1** *The standard greedy algorithm achieves approximation ratio  $\approx 5.95$ .*

Before getting into the proof, we first prove two algebraic identities.

**Lemma 4.2**

$$\sum_{j=1}^n \left(\frac{i+1}{i}\right)^{j-1} = i\left(\frac{i+1}{i}\right)^n - i.$$

**Proof:** Note that the expression on the left-hand side is a geometric sum. Therefore, we have

$$\sum_{j=1}^n \left(\frac{i+1}{i}\right)^{j-1} = \frac{\left(\frac{i+1}{i}\right)^n - 1}{\frac{i+1}{i} - 1} = i\left(\frac{i+1}{i}\right)^n - i.$$

□

**Lemma 4.3**

$$\sum_{j=1}^n j\left(\frac{i+1}{i}\right)^{j-1} = ni^2\left(\frac{i+1}{i}\right)^{n+1} - (n+1)i^2\left(\frac{i+1}{i}\right)^n + i^2.$$

**Proof:** Consider the function  $f(x) = \sum_{j=1}^n x^j$  with  $x \neq 1$ , its derivative  $f'(x) = \sum_{j=1}^n jx^{j-1}$ . Since  $f(x)$  is a geometric sum and  $x \neq 1$ , we have

$$f(x) = \frac{x^{n+1} - 1}{x - 1}.$$

Taking derivatives on both sides we have

$$f'(x) = \frac{(n+1)x^n(x-1) - x^{n+1} + 1}{(x-1)^2} = \frac{nx^{n+1} - (n+1)x^n + 1}{(x-1)^2}.$$

Therefore, we have

$$\sum_{j=1}^n jx^{j-1} = \frac{nx^{n+1} - (n+1)x^n + 1}{(x-1)^2}.$$

Substituting  $x$  with  $\frac{i+1}{i}$ , we have

$$\sum_{j=1}^n j\left(\frac{i+1}{i}\right)^{j-1} = \frac{n\left(\frac{i+1}{i}\right)^{n+1} - (n+1)\left(\frac{i+1}{i}\right)^n + 1}{\left(\frac{i+1}{i} - 1\right)^2} = ni^2\left(\frac{i+1}{i}\right)^{n+1} - (n+1)i^2\left(\frac{i+1}{i}\right)^n + i^2.$$

□

Now we proceed to the proof to Theorem 4.1.

**Proof:** Let  $S_i$  be the greedy solution after the  $i^{\text{th}}$  iteration; i.e.,  $|S_i| = i$ . Let  $O$  be an optimal solution, and let  $C_i = O \setminus S_i$ . Let  $m_i = |C_i|$ , and  $C_i = \{c_1, c_2, \dots, c_{m_i}\}$ . By the weakly submodularity definition, we get the following  $m_i$  inequalities for each  $0 < i < p$ :

$$\begin{aligned}
(i + m_i - 1)f(S_i \cup \{c_1\}) + (i + 1)f(S_i \cup \{c_2, \dots, c_{m_i}\}) &\geq (i)f(S_i \cup \{c_1, \dots, c_{m_i}\}) + (i + m_i)f(S_i) \\
(i + m_i - 2)f(S_i \cup \{c_2\}) + (i + 1)f(S_i \cup \{c_3, \dots, c_{m_i}\}) &\geq (i)f(S_i \cup \{c_2, \dots, c_{m_i}\}) + (i + m_i - 1)f(S_i) \\
&\vdots \\
(i + 1)f(S_i \cup \{c_{m_i-1}\}) + (i + 1)f(S_i \cup \{c_{m_i}\}) &\geq (i)f(S_i \cup \{c_{m_i-1}, c_{m_i}\}) + (i + 2)f(S_i) \\
(i)f(S_i \cup \{c_{m_i}\}) + (i + 1)f(S_i) &\geq (i)f(S_i \cup \{c_{m_i}\}) + (i + 1)f(S_i).
\end{aligned}$$

Multiplying the  $j^{\text{th}}$  inequality by  $(\frac{i+1}{i})^{j-1}$ , and summing all of them up (noting that the second term of the left hand side of the  $j^{\text{th}}$  inequality then cancels the first term of the  $j + 1^{\text{st}}$  inequality), we have

$$\begin{aligned}
\sum_{j=1}^{m_i} (i + m_i - j)(\frac{i+1}{i})^{j-1}f(S_i \cup \{c_j\}) + (i + 1)(\frac{i+1}{i})^{m_i-1}f(S_i) \\
\geq (i)f(S_i \cup \{c_1, \dots, c_{m_i}\}) + \sum_{j=1}^{m_i} (i + m_i - j + 1)(\frac{i+1}{i})^{j-1}f(S_i).
\end{aligned}$$

By monotonicity, we have  $f(S_i \cup \{c_1, \dots, c_{m_i}\}) \geq f(O)$ . Rearranging the inequality,

$$\sum_{j=1}^{m_i} (i + m_i - j)(\frac{i+1}{i})^{j-1}f(S_i \cup \{c_j\}) \geq (i)f(O) + \sum_{j=1}^{m_i-1} (i + m_i - j + 1)(\frac{i+1}{i})^{j-1}f(S_i).$$

By the greedy selection rule, we know that  $f(S_{i+1}) \geq f(S_i \cup \{c_j\})$  for any  $1 \leq j \leq m_i$ , therefore we have

$$\sum_{j=1}^{m_i} (i + m_i - j)(\frac{i+1}{i})^{j-1}f(S_{i+1}) \geq (i)f(O) + \sum_{j=1}^{m_i-1} (i + m_i - j + 1)(\frac{i+1}{i})^{j-1}f(S_i).$$

For the ease of notation, we let

$$a_i = \sum_{j=1}^{m_i} (i + m_i - j)(\frac{i+1}{i})^{j-1} \quad b_i = \sum_{j=1}^{m_i-1} (i + m_i - j + 1)(\frac{i+1}{i})^{j-1}$$

so that we have  $a_i f(S_{i+1}) - b_i f(S_i) \geq (i)f(O)$

We first simplify  $a_i$  and  $b_i$ .

$$\begin{aligned}
a_i &= \sum_{j=1}^{m_i} (i + m_i - j)(\frac{i+1}{i})^{j-1} \\
&= \sum_{j=1}^{m_i} (i + m_i)(\frac{i+1}{i})^{j-1} - \sum_{j=1}^{m_i} j(\frac{i+1}{i})^{j-1}.
\end{aligned}$$

By Lemma 4.2 and 4.3, we have

$$\begin{aligned}
a_i &= (i + m_i)[i(\frac{i+1}{i})^{m_i} - i] - m_i i^2 (\frac{i+1}{i})^{m_i+1} + (m_i + 1) i^2 (\frac{i+1}{i})^{m_i} - i^2 \\
&= [i^2 + i m_i - m_i(i^2 + i) + (m_i + 1) i^2] (\frac{i+1}{i})^{m_i} - 2i^2 - i m_i \\
&= 2i^2 (\frac{i+1}{i})^{m_i} - 2i^2 - i m_i.
\end{aligned}$$

Similarly, we have

$$\begin{aligned}
b_i &= \sum_{j=1}^{m_i-1} (i + m_i - j + 1) \left(\frac{i+1}{i}\right)^{j-1} \\
&= \sum_{j=1}^{m_i-1} (i + m_i + 1) \left(\frac{i+1}{i}\right)^{j-1} - \sum_{j=1}^{m_i-1} j \left(\frac{i+1}{i}\right)^{j-1} \\
&= (i + m_i + 1) [i \left(\frac{i+1}{i}\right)^{m_i-1} - i] - (m_i - 1) i^2 \left(\frac{i+1}{i}\right)^{m_i} + m_i i^2 \left(\frac{i+1}{i}\right)^{m_i-1} - i^2 \\
&= [i^2 + i m_i + i - (m_i - 1)(i^2 + i) + m_i i^2] \left(\frac{i+1}{i}\right)^{m_i-1} - 2i^2 - i m_i - i \\
&= 2i(i+1) \left(\frac{i+1}{i}\right)^{m_i-1} - 2i^2 - i m_i - i \\
&= 2i^2 \left(\frac{i+1}{i}\right)^{m_i} - 2i^2 - i m_i - i.
\end{aligned}$$

Now let

$$a_i^* = \sum_{j=1}^p (i + p - j) \left(\frac{i+1}{i}\right)^{j-1} \quad b_i^* = \sum_{j=1}^{p-1} (i + p - j + 1) \left(\frac{i+1}{i}\right)^{j-1}$$

The simplification of  $a_i$  and  $b_i$  makes it clear that  $a_i - b_i = i$  for any value of  $m_i$ . Since  $a_i^*$  (resp.  $b_i^*$ ) can be thought of as  $a_i$  (resp.  $b_i$ ) with  $m_i = p$ , we have

$$a_i^* - a_i = b_i^* - b_i \geq 0$$

Therefore,

$$a_i^* f(S_{i+1}) - b_i^* f(S_i) = a_i f(S_{i+1}) - b_i f(S_i) + (a_i^* - a_i)[f(S_{i+1}) - f(S_i)].$$

Since  $f(\cdot)$  is monotone, we have  $f(S_{i+1}) - f(S_i) \geq 0$ . Therefore,

$$a_i^* f(S_{i+1}) - b_i^* f(S_i) \geq a_i f(S_{i+1}) - b_i f(S_i) \geq i f(O).$$

Then we have the following set of inequalities:

$$\begin{aligned}
a_1^* f(S_2) &\geq 1 f(O) + b_1^* f(S_1) \\
a_2^* f(S_3) &\geq 2 f(O) + b_2^* f(S_2) \\
&\vdots \\
a_{p-2}^* f(S_{p-1}) &\geq (p-2) f(O) + b_{p-2}^* f(S_{p-2}) \\
a_{p-1}^* f(S_p) &\geq (p-1) f(O) + b_{p-1}^* f(S_{p-1}).
\end{aligned}$$

Multiplying the  $i^{\text{th}}$  inequality by  $\frac{\prod_{j=1}^{i-1} a_j^*}{\prod_{j=2}^i b_j^*}$ , summing all of them up and ignoring the term  $b_1^* f(S_1)$ ,

$$\frac{\prod_{j=1}^{p-1} a_j^*}{\prod_{j=2}^{p-1} b_j^*} f(S_p) \geq \sum_{i=1}^{p-1} \frac{i \prod_{j=1}^{i-1} a_j^*}{\prod_{j=2}^i b_j^*} f(O).$$

Therefore the approximation ratio

$$\frac{f(O)}{f(S_p)} \leq \frac{\frac{\prod_{j=1}^{p-1} a_j^*}{\prod_{j=2}^{p-1} b_j^*}}{\sum_{i=1}^{p-1} \frac{i \prod_{j=1}^{i-1} a_j^*}{\prod_{j=i}^{p-1} b_j^*}} = \left( \sum_{i=1}^{p-1} \frac{i \prod_{j=i+1}^{p-1} b_j^*}{\prod_{j=i}^{p-1} a_j^*} \right)^{-1} = \left( \sum_{i=1}^{p-1} \left[ \frac{i}{a_i^*} \cdot \prod_{j=i+1}^{p-1} \frac{b_j^*}{a_j^*} \right] \right)^{-1}.$$

Note that the approximation ratio is simply a function of  $p$ . In particular, the approximation ratio is 3.74 when  $p = 10$  and approximation ratio is 5.62 when  $p = 100$ . Computer evaluations suggest that the approximation ratio converges to 5.95 as  $p$  tends to  $\infty$ .  $\square$

In terms of hardness of approximation, assuming  $P \neq NP$ , Feige [7] proved that the max coverage problem (an example of monotone submodular maximization subject to a cardinality constraint) is known to be hard to approximate to a factor better than  $\frac{e}{e-1} - \epsilon$ . The problem of maximizing the sum of metric distances subject to a cardinality constraint has been called the *max-sum dispersion problem*. The max-sum dispersion problem is known to be NP-hard by an easy reduction from Max-Clique, and as noted by Alon [1], there is evidence that the problem is hard to compute in polynomial time with approximation  $2 - \epsilon$  for any  $\epsilon > 0$  when  $p = n^r$  for  $1/3 \leq r < 1$ . (See the discussion in Section 3 of [4].)

## 5 Weakly Submodular Function Maximization Subject to an Arbitrary Matroid Constraint

It is natural to consider a general matroid constraint for the problem of weakly submodular function maximization. For this more general problem, the greedy algorithm in the previous section no longer achieves any constant approximation ratio. See the example presented in the Appendix of [4]. Following the result for max-sum diversification subject to a matroid constraint in [5], we will analyze the following oblivious local search algorithm:

### WEAKLY SUBMODULAR FUNCTION MAXIMIZATION WITH A MATROID CONSTRAINT

- 1: Let  $S$  be a basis of  $\mathcal{M}$
- 2: **while** exists  $u \in U \setminus S$  and  $v \in S$  such that  $S \cup \{u\} \setminus \{v\} \in \mathcal{F}$  and  $f(S \cup \{u\} \setminus \{v\}) > f(S)$  **do**
- 3:      $S = S \cup \{u\} \setminus \{v\}$
- 4: **end while**
- 5: return  $S$

The following lemma on the exchange property of matroid bases was first stated in [6].

**Lemma 5.1 (Brualdi [6])** *For any two sets  $X, Y \in \mathcal{F}$  with  $|X| = |Y|$ , there is a bijective mapping  $g : X \rightarrow Y$  such that  $X \cup \{g(x)\} \setminus \{x\} \in \mathcal{F}$  for any  $x \in X$ .*

Before we prove the theorem, we need to prove several lemmas. Let  $O$  be the optimal solution, and  $S$ , the solution at the end of the local search algorithm. Let  $s$  be the size of a basis; let  $A = O \cap S$ ,  $B = S \setminus A$  and  $C = O \setminus A$ . By Lemma 5.1, there is a bijective mapping  $g : B \rightarrow C$  such that  $S \cup \{b\} \setminus \{g(b)\} \in \mathcal{F}$  for any  $b \in B$ . Let  $B = \{b_1, b_2, \dots, b_t\}$ , and let  $c_i = g(b_i)$  for all  $i = 1, \dots, t$ . We reorder  $b_1, b_2, \dots, b_t$  in different ways. Let  $b'_1, b'_2, \dots, b'_t$  be an ordering such that the corresponding  $c'_1, c'_2, \dots, c'_t$  maximizes the sum  $\sum_{i=1}^t (s-i)(\frac{s+1}{s})^{i-1} f(S \cup \{c'_i\})$ ; and let  $b''_1, b''_2, \dots, b''_t$  be an ordering such that the corresponding  $c''_1, c''_2, \dots, c''_t$  minimizes the sum

$$\sum_{i=1}^t (s+t-i)(\frac{s+1}{s})^{i-1} f(S \cup \{c''_i\}).$$

**Lemma 5.2** *Given three non-increasing non-negative sequences:*

$$\begin{aligned}\alpha_1 &\geq \alpha_2 \geq \cdots \geq \alpha_n \geq 0, \\ \beta_1 &\geq \beta_2 \geq \cdots \geq \beta_n \geq 0, \\ x_1 &\geq x_2 \geq \cdots \geq x_n \geq 0.\end{aligned}$$

*Then we have*

$$\sum_{i=1}^n \alpha_i x_i \sum_{i=1}^n \beta_i \geq \sum_{i=1}^n \beta_i x_{n+1-i} \sum_{i=1}^n \alpha_i.$$

**Proof:** Consider the following:

$$\begin{aligned}n \sum_{i=1}^n \alpha_i x_i &= n\alpha_1 x_1 + n\alpha_2 x_2 + \cdots + n\alpha_n x_n \\ &= \sum_{i=1}^n \alpha_i x_1 + (n\alpha_1 - \sum_{i=1}^n \alpha_i) x_1 + n\alpha_2 x_2 + \cdots + n\alpha_n x_n \\ &\geq \sum_{i=1}^n \alpha_i x_1 + (n\alpha_1 + n\alpha_2 - \sum_{i=1}^n \alpha_i) x_2 + \cdots + n\alpha_n x_n \\ &= \sum_{i=1}^n \alpha_i x_1 + \sum_{i=1}^n \alpha_i x_2 + (n\alpha_1 + n\alpha_2 - 2 \sum_{i=1}^n \alpha_i) x_2 + \cdots + n\alpha_n x_n \\ &\quad \vdots \\ &\geq \sum_{i=1}^n \alpha_i x_1 + \sum_{i=1}^n \alpha_i x_2 + \cdots + \sum_{i=1}^n \alpha_i x_n + (n\alpha_1 + n\alpha_2 + \cdots + n\alpha_n - n \sum_{i=1}^n \alpha_i) x_n \\ &= \sum_{i=1}^n \alpha_i \sum_{i=1}^n x_i\end{aligned}$$

Similarly, we have

$$\begin{aligned}n \sum_{i=1}^n \beta_i x_{n+1-i} &= n\beta_1 x_n + n\beta_2 x_{n-1} + \cdots + n\beta_n x_1 \\ &= \sum_{i=1}^n \beta_i x_n + (n\beta_1 - \sum_{i=1}^n \beta_i) x_n + n\beta_2 x_{n-1} + \cdots + n\beta_n x_1 \\ &\leq \sum_{i=1}^n \beta_i x_n + (n\beta_1 + n\beta_2 - \sum_{i=1}^n \beta_i) x_{n-1} + \cdots + n\beta_n x_1 \\ &= \sum_{i=1}^n \beta_i x_n + \sum_{i=1}^n \beta_i x_{n-1} + (n\beta_1 + n\beta_2 - 2 \sum_{i=1}^n \beta_i) x_{n-1} + \cdots + n\beta_n x_1 \\ &\quad \vdots \\ &\leq \sum_{i=1}^n \beta_i x_n + \sum_{i=1}^n \beta_i x_{n-1} + \cdots + \sum_{i=1}^n \beta_i x_1 + (n\alpha_1 + n\alpha_2 + \cdots + n\alpha_n - n \sum_{i=1}^n \alpha_i) x_1 \\ &= \sum_{i=1}^n \beta_i \sum_{i=1}^n x_i\end{aligned}$$

Therefore the lemma follows.  $\square$

**Lemma 5.3**

$$\begin{aligned} & \sum_{i=1}^t (s-i) \left(\frac{s+1}{s}\right)^{i-1} f(S \cup \{c'_i\}) \\ & \leq sf(S) + \sum_{i=1}^t (s+1-i) \left(\frac{s+1}{s}\right)^{i-1} f(S \cup \{c'_i\} \setminus \{b'_i\}) - (s+1) \left(\frac{s+1}{s}\right)^{t-1} f(S \setminus \{b'_1, \dots, b'_t\}). \end{aligned}$$

**Proof:** By the definition of weakly submodular, we have

$$\begin{aligned} sf(S) + sf(S \cup \{c'_1\} \setminus \{b'_1\}) & \geq (s-1)f(S \cup \{c'_1\}) + (s+1)f(S \setminus \{b'_1\}) \\ sf(S \setminus \{b'_1\}) + (s-1)f(S \cup \{c'_2\} \setminus \{b'_2\}) & \geq (s-2)f(S \cup \{c'_2\}) + (s+1)f(S \setminus \{b'_1, b'_2\}) \\ & \vdots \\ sf(S \setminus \{b'_1, \dots, b'_{t-1}\}) + (s-t+1)f(S \cup \{c'_t\} \setminus \{b'_t\}) & \geq (s-t)f(S \cup \{c'_t\}) + (s+1)f(S \setminus \{b'_1, \dots, b'_t\}) \end{aligned}$$

Multiplying the  $i^{\text{th}}$  inequality by  $(\frac{s+1}{s})^{i-1}$ , and summing all of them up to get

$$\begin{aligned} & sf(S) + \sum_{i=1}^t (s+1-i) \left(\frac{s+1}{s}\right)^{i-1} f(S \cup \{c'_i\} \setminus \{b'_i\}) \\ & \geq \sum_{i=1}^t (s-i) \left(\frac{s+1}{s}\right)^{i-1} f(S \cup \{c'_i\}) + (s+1) \left(\frac{s+1}{s}\right)^{t-1} f(S \setminus \{b'_1, \dots, b'_t\}). \end{aligned}$$

After rearranging the inequality, we get

$$\begin{aligned} & \sum_{i=1}^t (s-i) \left(\frac{s+1}{s}\right)^{i-1} f(S \cup \{c'_i\}) \\ & \leq sf(S) + \sum_{i=1}^t (s+1-i) \left(\frac{s+1}{s}\right)^{i-1} f(S \cup \{c'_i\} \setminus \{b'_i\}) - (s+1) \left(\frac{s+1}{s}\right)^{t-1} f(S \setminus \{b'_1, \dots, b'_t\}). \end{aligned}$$

□

**Lemma 5.4**

$$\begin{aligned} & \sum_{i=1}^t (s+t-i) \left(\frac{s+1}{s}\right)^{i-1} f(S \cup \{c''_i\}) - \sum_{i=1}^t (s+t+1-i) \left(\frac{s+1}{s}\right)^{i-1} f(S) \\ & \geq sf(S \cup \{c''_1, \dots, c''_t\}) - (s+1) \left(\frac{s+1}{s}\right)^{t-1} f(S) \end{aligned}$$

**Proof:** By the definition of weakly submodular, we have

$$\begin{aligned} (s+t-1)f(S \cup \{c''_1\}) + (s+1)f(S \cup \{c''_2, \dots, c''_{m_i}\}) & \geq sf(S \cup \{c''_1, \dots, c''_{m_i}\}) + (s+t)f(S) \\ & \vdots \\ (s+1)f(S \cup \{c''_{t-1}\}) + (s+1)f(S \cup \{c''_t\}) & \geq sf(S \cup \{c''_{t-1}, c''_t\}) + (s+2)f(S) \\ sf(S \cup \{c''_t\}) + (s+1)f(S) & \geq sf(S \cup \{c''_t\}) + (s+1)f(S). \end{aligned}$$

Multiplying the  $i^{\text{th}}$  inequality by  $(\frac{s+1}{s})^{i-1}$ , and summing all of them up, we have

$$\begin{aligned} \sum_{i=1}^t (s+t-i)(\frac{s+1}{s})^{i-1} f(S \cup \{c''_i\}) + (s+1)(\frac{s+1}{s})^{t-1} f(S) \\ \geq sf(S \cup \{c''_1, \dots, c''_t\}) + \sum_{i=1}^t (s+t+1-i)(\frac{s+1}{s})^{i-1} f(S). \end{aligned}$$

Therefore, we have

$$\begin{aligned} \sum_{i=1}^t (s+t-i)(\frac{s+1}{s})^{i-1} f(S \cup \{c''_i\}) \\ \geq sf(S \cup \{c''_1, \dots, c''_t\}) + \sum_{i=1}^t (s+t+1-i)(\frac{s+1}{s})^{i-1} f(S) - (s+1)(\frac{s+1}{s})^{t-1} f(S). \end{aligned}$$

□

Let

$$\begin{aligned} W &= \sum_{i=1}^t (s-i)(\frac{s+1}{s})^{i-1}, & X &= \sum_{i=1}^t (s+1-i)(\frac{s+1}{s})^{i-1}, \\ Y &= \sum_{i=1}^t (s+t-i)(\frac{s+1}{s})^{i-1}, & Z &= \sum_{i=1}^t (s+t+1-i)(\frac{s+1}{s})^{i-1}. \end{aligned}$$

**Lemma 5.5**

$$C \sum_{i=1}^t (s-i)(\frac{s+1}{s})^{i-1} f(S \cup \{c'_i\}) \geq A \sum_{i=1}^t (s+t-i)(\frac{s+1}{s})^{i-1} f(S \cup \{c''_i\}).$$

**Proof:** This is immediate by Lemma 5.2

□

**Theorem 5.6** *Let  $s$  be the size of a basis, the local search algorithm achieves an approximation ratio bounded by 14.5 for an arbitrary  $s$ , approximately 10.88 when  $s = 6$ . The ratio converges to 10.22 as  $s$  tends to  $\infty$ .*

**Proof:** Since  $S$  is a locally optimal solution, we have

$$f(S) \geq f(S \cup \{c'_i\} \setminus \{b'_i\}).$$

Since  $f(S \setminus \{b'_1, \dots, b'_t\}) \geq 0$ , by Lemma 5.3, we have

$$\sum_{i=1}^t (s-i)(\frac{s+1}{s})^{i-1} f(S \cup \{c'_i\}) \leq sf(S) + \sum_{i=1}^t (s+1-i)(\frac{s+1}{s})^{i-1} f(S).$$

Therefore,

$$\sum_{i=1}^t (s-i)(\frac{s+1}{s})^{i-1} f(S \cup \{c'_i\}) \leq (s+X) f(S).$$

On the other hand, we have  $O \subseteq S \cup \{c_1'', \dots, c_t''\}$ , by monotonicity, we have  $f(O) \leq f(S \cup \{c_1'', \dots, c_t''\})$ . By Lemma 5.4, we have

$$\sum_{i=1}^t (s+t-i) \left(\frac{s+1}{s}\right)^{i-1} f(S \cup \{c_i''\}) \geq s f(O) + [Z - (s+1) \left(\frac{s+1}{s}\right)^{t-1}] f(S).$$

Lemma 5.2, we have

$$Y \sum_{i=1}^t (s-i) \left(\frac{s+1}{s}\right)^{i-1} f(S \cup \{c_i'\}) \geq W \sum_{i=1}^t (s+t-i) \left(\frac{s+1}{s}\right)^{i-1} f(S \cup \{c_i''\}).$$

Therefore

$$Y(s+X)f(S) \geq Ws f(O) + X[Z - (s+1) \left(\frac{s+1}{s}\right)^{t-1}] f(S)$$

Hence the approximation ratio:

$$\frac{f(O)}{f(S)} \leq \frac{YX - WZ + Ys + W(s+1) \left(\frac{s+1}{s}\right)^{t-1}}{Ws} = \frac{YX - WZ + Ys}{Ws} + \left(\frac{s+1}{s}\right)^t.$$

Simplifying the notation, we have

$$\frac{f(O)}{f(S)} \leq \frac{\sum_{i=1}^t (s^2 + st + ti - si) \left(\frac{s+1}{s}\right)^{i-1} + \sum_{i=t+1}^{2t-1} t(2t-i) \left(\frac{s+1}{s}\right)^{i-1}}{\sum_{i=1}^t s(s-i) \left(\frac{s+1}{s}\right)^{i-1}} + \left(\frac{s+1}{s}\right)^t.$$

Using Lemma 4.2 and 4.3 to simply it further, we have

$$\frac{f(O)}{f(S)} \leq \frac{2s \left(\frac{s+1}{s}\right)^{2t} - 2t \left(\frac{s+1}{s}\right)^t - 2s}{(2s-t) \left(\frac{s+1}{s}\right)^t - 2s}.$$

Let  $x = \left(\frac{s+1}{s}\right)^s$  and  $r = \frac{t}{s}$ , we study the continuous version of the above function

$$g(x, r) = \frac{2x^{2r} - 2rx^r - 2}{(2-r)x^r - 2}.$$

Since  $S$  is a local optimum with respect to the swapping of any single element and by the definition of  $x, s$  and  $t$ , we have  $2 \leq t \leq s$  and hence  $2.25 \leq x \leq e$  and  $0 < r \leq 1$ . Our goal then is to establish an upper bound on  $g(x, r)$  for  $2.25 \leq x \leq e$  and  $0 < r \leq 1$ . We will think of  $g(x, r)$  as implicitly defining  $x$  as a function of  $r$  at points where  $g(x, r)$  can possibly take on a maximum value, namely when  $\frac{\partial g(x, r)}{\partial x} = 0$  and at the boundary points for  $x$ .

Note that since  $x \geq 2.25$ ,

$$x > \left(\frac{2}{2-r}\right)^{\frac{1}{r}},$$

for all  $0 < r \leq 1$ . Therefore, we have  $(2-r)x^r - 2 > 0$  for given  $x$  and  $r$ . It is easy to verify that function  $g(x, r)$  is continuous and differentiable. For any fixed  $r$ , the function has two boundary points at  $x = 2.25$  and  $x = e$ , and taking partial derivative with respect to  $x$ , we have

$$\frac{\partial g(x, r)}{\partial x} = \frac{2rx^{r-1}(x^r - 1)[(2-r)x^r - (2+r)]}{[(2-r)x^r - 2]^2}.$$

Therefore the only point where the partial derivative equals to zero is

$$x^* = \left(\frac{2+r}{2-r}\right)^{\frac{1}{r}}.$$

Plugging this into the original expression for  $g(x, r)$ , we have

$$g(x^*, r) = \frac{2r^2 + 8}{(r-2)^2}.$$

The function  $g(x^*, r)$  is monotonically increasing with respect to  $r \in (0, 1]$  and it has a maximum value of 10 when  $r = 1$ .

Now it only remains to check the two boundary points  $x = 2.25$  and  $x = e$ . Note that these are fixed values. We now fix  $x$ , and take partial derivative with respect to  $r$ :

$$\frac{\partial g(x, r)}{\partial r} = \frac{2x^r(r-1)[(2\ln x - r\ln x + 1)x^r - (2\ln x + r\ln x + 1)]}{[(2-r)x^r - 2]^2}.$$

Since  $x^r > 0$ ,  $x^r - 1 > 0$  and  $[(2-r)x^r - 2]^2 > 0$ . If we can show that

$$(2\ln x - r\ln x + 1)x^r - (2\ln x + r\ln x + 1) > 0$$

then the function after fixing  $x$  is monotonically increasing with respect to  $r$ . We use the Taylor expansion of  $x^r$  at  $x = 0$ .

$$x^r > 1 + r\ln x + \frac{1}{2}r^2\ln^2 x.$$

Therefore,

$$(2\ln x - r\ln x + 1)x^r - (2\ln x + r\ln x + 1) > r\ln x(2\ln x + r\ln^2 x - \frac{1}{2}r^2\ln^2 x - \frac{1}{2}r\ln x - 1).$$

Note that we only need to check for the case when  $x = e$  and  $x = 2.25$ .

1. Case  $x = e$ :

$$2\ln x + r\ln^2 x - \frac{1}{2}r^2\ln^2 x - \frac{1}{2}r\ln x - 1 = 1 + \frac{1}{2}r - \frac{1}{2}r^2 > 0.$$

2. Case  $x = 2.25$ :

$$2\ln x + r\ln^2 x - \frac{1}{2}r^2\ln^2 x - \frac{1}{2}r\ln x - 1 > 0.6 + 0.6r - 0.5r - 0.4r^2 > 0.$$

Therefore  $(2\ln x - r\ln x + 1)x^r - (2\ln x + r\ln x + 1) > 0$ , and hence  $\frac{\partial g(x, r)}{\partial r} > 0$  for  $x = 2.25$  and  $x = e$ . Therefore the maximum is obtained when  $r = 1$ . Plug  $r = 1$  into the original formula, we have

$$g(x, 1) = \frac{2x^2 - 2x - 2}{x - 2}.$$

Evaluate it for  $x = e$  and  $x = 2.25$ , we have  $g(e, 1) = 10.22$  and  $g(2.25, 1) = 14.5$ . This completes the proof.  $\square$

## 6 Conclusion and Open Problem

Motivated by the max-sum diversification problem we are led to study a generalization of monotone submodular functions that we call weakly-submodular functions. This class includes the supermodular max-sum dispersion problem.

There are several open problems that remain. First, similar to the result for an arbitrary matroid constraint, we would like to have a proof of the convergence of the approximation bound for the cardinality constraint. Another immediate open problem is to close the gap between the upper and lower bounds we know for approximating an arbitrary weakly submodular function subject to cardinality or matroid constraints. It would also be of interest to consider an approximation for maximizing a weakly submodular function subject to a knapsack constraint. In addition, we ask what other possible extensions of submodular functions can be defined so as to include supermodular functions and yet be amenable to simple approximation algorithms. Finally, we would like to know if there is an analogue of the marginal decreasing property that characterizes submodular functions.

**Acknowledgment** This research is supported by the Natural Sciences and Engineering Research Council of Canada and the University of Toronto, Department of Computer Science.

## References

- [1] N. Alon. Personal communication.
- [2] N. Alon, S. Arora, R. Manoikaran, D. Moshkovitz, and O. Weinstein. Inapproximability of densest  $\kappa$ -subgraph from average case hardness. *Unpublished manuscript*, 2011.
- [3] N. Bansal, K. Jain, A. Kazeykina, and J. Naor. Approximation algorithms for diversified search ranking. In *Proceedings of the 37th international colloquium conference on Automata, languages and programming: Part II*, ICALP’10, pages 273–284, Berlin, Heidelberg, 2010. Springer-Verlag.
- [4] A. Borodin, A. Jain, H. C. Lee, and Y. Ye. Max-sum diversification, monotone submodular functions and dynamic updates. *arXiv and Journal version of 2012 PODS conference paper*, 2014.
- [5] A. Borodin, H. C. Lee, and Y. Ye. Max-sum diversification, monotone submodular functions and dynamic updates. In *PODS*, pages 155–166, 2012.
- [6] R. A. Brualdi. Comments on bases in dependence structures. *Bulletin of the Australian Mathematical Society*, 1:161–167, 8 1969.
- [7] U. Feige. A threshold of  $\ln n$  for approximating set cover. *J. ACM*, 45(4):634–652, July 1998.
- [8] S. Gollapudi and A. Sharma. An axiomatic approach for result diversification. In *Proceedings of the 18th international conference on World wide web*, WWW ’09, pages 381–390, New York, NY, USA, 2009. ACM.
- [9] D. Kempe, J. Kleinberg, and E. Tardos. Maximizing the spread of influence through a social network. In *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining*, KDD ’03, pages 137–146, New York, NY, USA, 2003. ACM.

- [10] J. Kleinberg, C. Papadimitriou, and P. Raghavan. Segmentation problems. In *Proceedings of the thirtieth annual ACM symposium on Theory of computing*, STOC '98, pages 473–482, New York, NY, USA, 1998. ACM.
- [11] M. R. Korupolu, C. Plaxton, and R. Rajaraman. Analysis of a local search heuristic for facility location problems. *Journal of Algorithms*, 37(1):146–188, 2000.
- [12] H. Lin and J. Bilmes. A class of submodular functions for document summarization. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies - Volume 1*, HLT '11, pages 510–520, Stroudsburg, PA, USA, 2011. Association for Computational Linguistics.
- [13] G. Nemhauser, L. Wolsey, and M. Fisher. An analysis of the approximations for maximizing submodular set functions. *Mathematical Programming*, 1978.